SCIENTIFIC REPORTS

OPEN

Phylogenomics illuminates the backbone of the Myriapoda Tree of Life and reconciles morphological and molecular phylogenies

Received: 21 July 2017
Accepted: 14 December 2017
Published online: 08 January 2018

Rosa Fernández[1,3], Gregory D. Edgecombe[2] & Gonzalo Giribet[1]

The interrelationships of the four classes of Myriapoda have been an unresolved question in arthropod phylogenetics and an example of conflict between morphology and molecules. Morphology and development provide compelling support for Diplopoda (millipedes) and Pauropoda being closest relatives, and moderate support for Symphyla being more closely related to the diplopod-pauropod group than any of them are to Chilopoda (centipedes). In contrast, several molecular datasets have contradicted the Diplopoda–Pauropoda grouping (named Dignatha), often recovering a Symphyla–Pauropoda group (named Edafopoda). Here we present the first transcriptomic data including a pauropod and both families of symphylans, allowing myriapod interrelationships to be inferred from phylogenomic data from representatives of all main lineages. Phylogenomic analyses consistently recovered Dignatha with strong support. Taxon removal experiments identified outgroup choice as a critical factor affecting myriapod interrelationships. Diversification of millipedes in the Ordovician and centipedes in the Silurian closely approximates fossil evidence whereas the deeper nodes of the myriapod tree date to various depths in the Cambrian-Early Ordovician, roughly coinciding with recent estimates of terrestrialisation in other arthropod lineages, including hexapods and arachnids.

The evolutionary interrelationships between and within major arthropod groups were subject to much instability in the early years of molecular phylogenetics. Some hypotheses to emerge from that era – such as crustacean paraphyly with respect to Hexapoda (including insects) – have stood the test of time, whereas others have fallen by the wayside. Controversial results were exposed to be artefacts of insufficient amounts of data, flawed analytical methods, or systematic error. In recent years, phylogenomic approaches drawing on vastly expanded gene and taxon coverage, combined with improved analytical approaches, have seen stable, well supported molecular hypotheses being recovered[1–5], and these have eliminated several instances of incongruence with morphological trees that were introduced in earlier molecular studies.

Transcriptome-based phylogenies drawing on hundreds or thousands of orthologues have assisted with phylogenetic analyses for the major groups of millipedes[6,7] and centipedes[7,8] but relationships between the four main myriapod groups have not been as rigorously tested (Fig. 1). A particular limitation is that only Sanger-sequenced data are available for pauropods—a group that lies at the crux of a molecular and morphological conflict within Myriapoda—probably due to their small size, difficulty in finding them, and cryptic behaviour. Analyses based on 62 nuclear protein coding genes[9,10] underpin a formal taxonomic proposal that Pauropoda is most closely related to Symphyla, a putative clade named Edafopoda[11]. This grouping, also recovered using nuclear ribosomal genes[12] and mitochondrial genomes[13], is highly unexpected from the perspective of morphology because an alternative grouping of Pauropoda and Diplopoda has been widely accepted for more than a century[14–16]. This hypothesis is named Dignatha (=Collifera), referring to the mandibles and first maxillae being the only functional mouthparts, with the postmaxillary segment being limbless, parts of it forming a tergite called the *collum* and not incorporated into the head. Other shared morphological characters include a first maxilla coalesced with a sternal

[1]Museum of Comparative Zoology & Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford St., 02138, Cambridge, MA, USA. [2]Department of Earth Sciences, The Natural History Museum, Cromwell Road, London, SW7 5BD, UK. [3]Present address: Bioinformatics & Genomics, Centre for Genomic Regulation, Carrer del Dr. Aiguader 88, 08003, Barcelona, Spain. Correspondence and requests for materials should be addressed to R.F. (email: rfernandezgarcia@g.harvard.edu)

Figure 1. The four main groups of myriapods. (**A**) *Otostigmus* (*Parotostigmus*) *pococki* (Northern Range, Trinidad, Trinidad and Tobago) (Chilopoda, Scolopendromorpha). (**B**) *Hanseniella* sp. (South Island, New Zealand) (Symphyla, Scutigerellidae). (**C**) *Pauropus huxleyi* (Massachusetts, USA) (Pauropoda, Tetramerocerata). and (**D**) *Platydesmus* sp. (La Selva, Costa Rica) (Diplopoda, Platydesmida). Pictures by Gonzalo Giribet.

intermaxillary plate, the *vas deferens* opening to the tips of conical penes between the second trunk leg pair, and the spiracles opening to a tracheal pouch that functions as an apodeme. Early post-embryonic development unites Dignatha based on a motionless pupoid stage immediately after hatching, followed by a hexapodous first free-living stage. The Dignatha hypothesis has also been supported by a few smaller molecular data sets[17], but it has been contradicted by Edafopoda in the analysis of larger data sets [e.g., ref.[11]].

In order to investigate the conflicting support for Dignatha versus Edafopoda – and therefore to shed light on the backbone of the Myriapoda Tree of Life – we present the first transcriptomic data set including a pauropod and both families of symphylans. The new data are evaluated in a phylogenomic context specifically designed to test these hypotheses. Furthermore, we expand on previous efforts to date the myriapod phylogenetic tree by coding a morphological character set for the same set of species as sampled transcriptomically as well as key fossil species for their preserved morphological characters in order to estimate the age of diversification within Myriapoda, particularly with reference to the likely timing of terrestrialisation.

Results and Discussion

Pauropoda is the sister group to Diplopoda. All the analyses (with the exception of two, in which Pauropoda was attracted to the outgroups, see below) recovered Pauropoda as the sister group to Diplopoda with high support (Fig. 2a,b). Notably, support for Dignatha is strong when the most complete taxonomic sampling of non-myriapod outgroups is used (Fig. 2a). Given the unanimity of support for Dignatha/Collifera in morphological studies, this stable, well-supported result reconciles classical morphological studies with molecules. In analyses based on more intensively sampled or closely related outgroups (discussed below), the sister group of Dignatha is Symphyla, together forming the traditional clade Progoneata. The only two analyses not recovering Dignatha (both maximum likelihood analyses not accounting for among-site rate heterogeneity and including the most distant outgroups) positioned Pauropoda at the base of the ingroup due to a long branch attraction artefact (LBA) (Fig. 2c,d). In fact, one of them even recovered Myriapoda as non-monophyletic (Fig. 2d), with the pauropod spuriously clustering within Pancrustacea, highlighting the potential of LBA in this data set. Edafopoda was not recovered in either of these two analyses, as symphylans fall as the sister group of Diplopoda + Chilopoda in both cases (although without strong support in one of the analyses; Fig. 2c). The attraction of symphylans and pauropods as Edafopoda (the only hypothesis exclusively based on molecular information) is therefore probably due to artefacts during phylogenetic reconstruction, as discussed below.

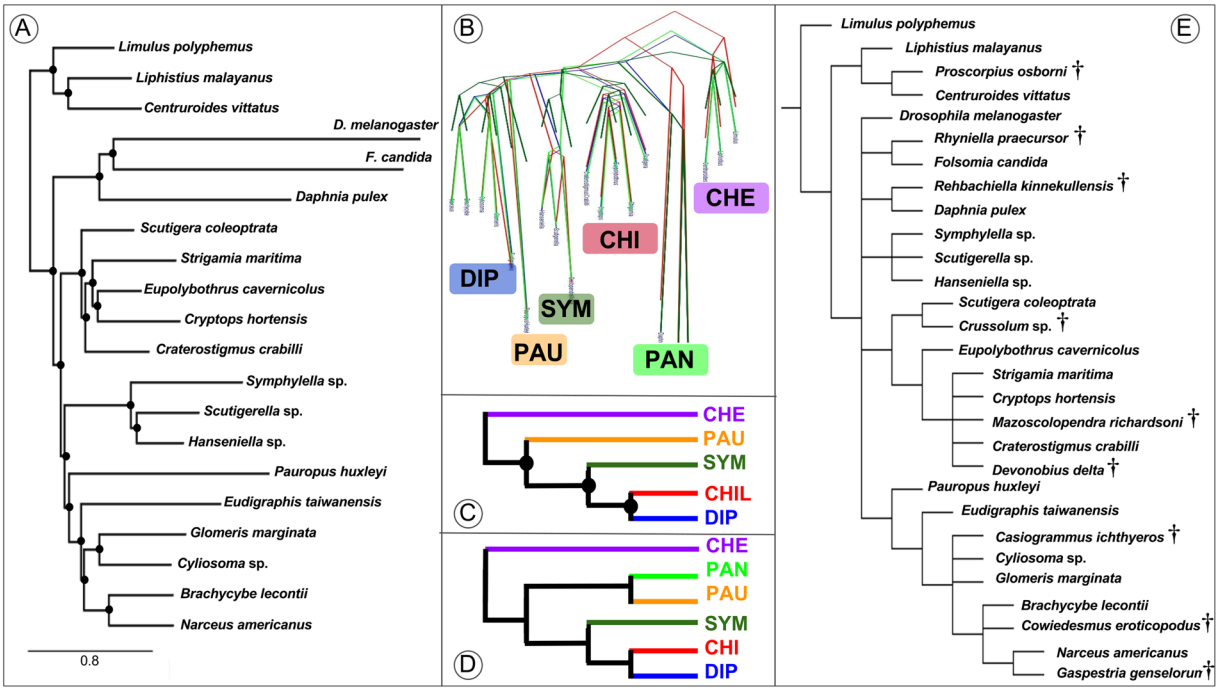

Figure 2. (**A**) Phylogenetic hypothesis of myriapod interrelationships (PhyloBayes, matrix 1). (**B**). DensiTree visualization of the four most congruent analyses (PhyloBayes, matrices 2 and 3; PhyML, matrix 3). (**C**,**D**). Main conflicting alternative hypotheses (**C**) PhyML, matrix 2; (**D**) PhyML, matrix 1. (**E**) Phylogenetic hypothesis of Myriapoda based on 187 morphological characters coded for both extant and extinct species (see Methods for further details); strict consensus of 105 trees of 257 steps; fossil taxa are identified with a dagger symbol. Black circles in nodes represent high support (>0.95 posterior probability, >90% bootstrap support). CHE: Chelicerata. PAN: Pancrustacea. CHI: Chilopoda. SYM: Symphyla. PAU: Pauropoda. DIP: Diplopoda. Colour codes for each clade are maintained in all figures.

Outgroup selection impacts on myriapod phylogeny. Despite Dignatha being recovered in most of our analyses, a result not often found in prior molecular studies, the interrelationships between it and the other two main groups of myriapods varied across analyses, being sensitive to outgroup choice. In the PhyloBayes analyses, matrices 1 and 3 recovered Progoneata, whereas in matrix 2 (in which only the more distantly allied chelicerates were selected as outgroups) symphylans appeared as sister group of the other Myriapoda instead of centipedes, as in the previous cases (Fig. 2a,b). The latter clade formed by these three groups was also recovered in the ML analyses (with or without strong support, though; Fig. 2c,d). This is not the first study in which this result was obtained: in Miyazawa *et al.*[18] symphylans were likewise recovered as sister group to all other myriapods, followed by pauropods as sister group to millipedes plus centipedes. However, that study was based on just three Sanger sequenced genes, and conflicts with other well resolved nodes in our phylogeny (e.g., Dignatha). The present study and Fernández *et al.*[7] suggest that outgroup selection is a major factor affecting phylogenetic reconstruction in myriapods. In addition, the latter study found that the most complete matrices were enriched in ribosomal proteins, and both factors strongly compromised the estimated relationships within the ingroup. In the present study, biases from ribosomal proteins were minimized by using a different orthology inference procedure, which ensures that only single copy genes are selected. In spite of this, it remains the case that the selection of only distant outgroups (chelicerates in this case) yields interrelationships of the myriapod classes that are less congruent with morphology than when closer and more comprehensively sampled outgroups are included. This study also highlights the importance of accounting for site-specific heterogeneity (through the CAT-GTR model of PhyloBayes) at least when taxon sampling is not dense for some of the groups, as even when only closer outgroups are included the long-branched pauropod is attracted to the equally long-branched Pancrustacea. The inclusion of more pauropods may alleviate this effect.

The timing of myriapod diversification. Diversification of Chilopoda (i.e., the basal split in the crown group) is dated to the Early Silurian (Fig. 3), not much earlier than the oldest fossil chilopods in the Late Silurian, these already being representatives of the chilopod crown group. Diversification of Diplopoda dates to the Middle Ordovician (autocorrelated rates)–earliest Silurian (uncorrelated rates). Though this is considerably older than the first millipede body fossils (from the Wenlock Series of the Silurian), it closely approximates the age of trace fossils that have been attributed to Diplopoda and especially compared to locomotion in Penicillata[19,20]. In contrast, deeper nodes associated with the divergences between myriapod classes are substantially older than available fossil data. No plausible total-group myriapod body fossils are known from the Cambrian, but as in previous studies dating Myriapoda[3,17], some deep splits are estimated to be of Cambrian age. Diversification of Dignatha is

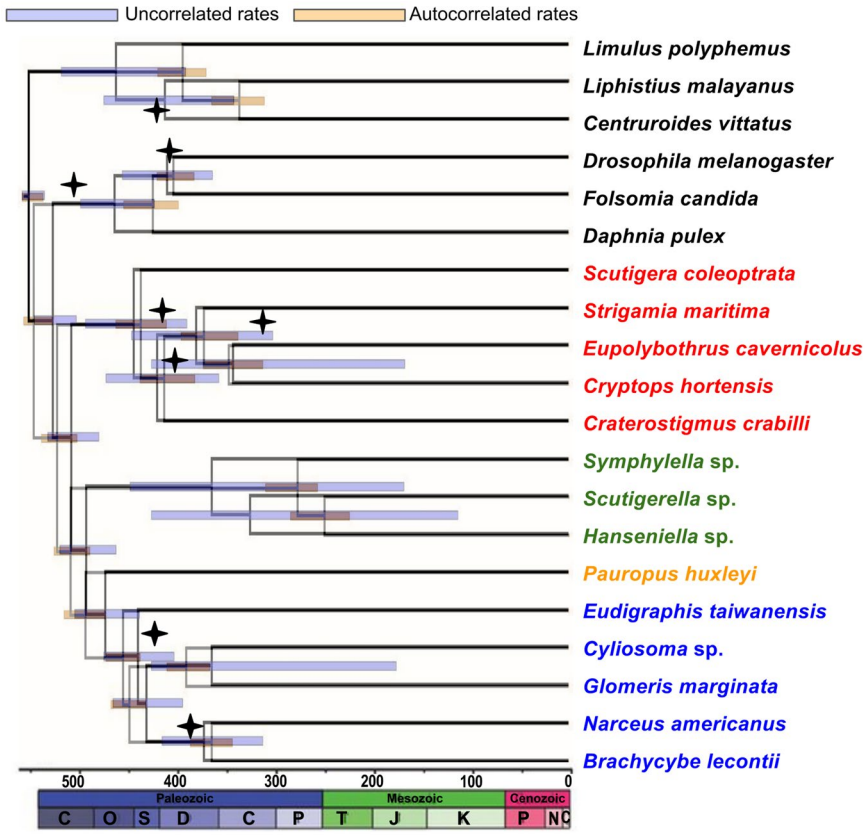

Figure 3. Chronogram of myriapod evolution based on matrix 1 (PhyloBayes analysis) with 95% highest posterior density (HPD) bars for the dating under the uncorrelated (blue) or autocorrelated (pink) model. Nodes that were calibrated with fossils are indicated with a diamond placed at the age of the fossil.

inferred to date to the latest Cambrian-Early Ordovician, Progoneata to the mid-late Cambrian, and Myriapoda to the early-middle Cambrian (auto- and uncorrelated rates, respectively). The shared terrestrial adaptations of all extant myriapods (e.g., tracheae, Malpighian tubules, uniramous trunk limbs) suggest that the common ancestor of each of these estimated Cambrian nodes was terrestrial, coinciding (although being slightly younger) with estimates of terrestrialisation for other arthropod lineages, including arachnids and hexapods[3,5]. Although the trace fossil record is consistent with amphibious arthropods by the mid Cambrian[21,22], and some such traces are potentially made by stem-group myriapods, current molecular estimates for early or middle Cambrian crown-group myriapods, earlier than the expected terrestrial flora, continue to pose an unanswered question in arthropod terrestrialisation.

Towards a fully-resolved Myriapoda Tree of Life. The branching pattern of the four main groups of myriapods has been one of the unresolved questions in arthropod phylogenetics, together with the interrelationships of the chelicerate orders and the exact sequence of crustaceans that led to the origins of hexapods. With the advent of phylogenomic methods, myriapod phylogeny has attracted attention during the last few years, with several studies devoted to shedding light on the interrelationships of the millipede[6] and centipede[8] orders, and more recently expanding taxon sampling to include most centipede families, most millipede orders and a couple of symphylans[7]. The different analyses of large data matrices combined in all these studies (as well as the current one) have allowed us to discern the main artefacts affecting phylogenetic reconstruction in this group of arthropods. The tree, its deep nodes congruent with traditional hypotheses based on morphology and development, can now be seen as a well-resolved backbone phylogeny with only a handful of untested placements, including the unsampled pauropod order Hexamerocerata, and the unexplored position of the diplopod orders Siphoniulida, Siphonophorida and Siphonocryptida. Some cases of incongruence between morphological and molecular data remain at shallower nodes, such as the interrelationships of the three orders of pentazonian millipedes[23], and the position of the centipede orders Craterostigmomorpha and Lithobiomorpha relative to each other and to Scolopendromorpha + Geophilomorpha[7].

Methods

Sample collection and molecular techniques. Fourteen species representing the four major groups of myriapods (Chilopoda, Diplopoda, Pauropoda and Symphyla) were included in this study. Building upon previous work[6,7], our sampling was designed to maximize representation of all groups, including all orders of centipedes, both families of symphylans, the main clades of millipedes, and pauropods. New sequence data were

Species	Data	Source	MCZ Voucher	SRA #
Chilopoda				
Scutigera coleoptrata	Transcriptome	Fernández *et al.*[8]	IZ-204015	SRR1158078
Craterostigmus crabilli	Transcriptome	Fernández *et al.*[7]	IZ-71256	SRR3232915
Cryptops hortensis	Transcriptome	Fernández *et al.*[8]	IZ-130583	SRR1153457
Strigamia maritima	Genome	Chipman *et al.*[35]	—	
Eupolybothrus cavernicolus	Transcriptome	Stoev *et al.*[36]	—	ERX311347
Symphyla				
Hanseniella sp.	Transcriptome	Fernández *et al.*[8]	IZ-133580	SRR6217953
Scutigerella sp.	Transcriptome	Fernández *et al.*[7]	IZ-46890	SRR3458649
Symphylella sp.	Transcriptome	Illumina HiSeq (this study)	IZ-141598	SRR6144316
Pauropoda				
Pauropus huxleyi	Transcriptome	Illumina HiSeq (this study)	IZ-141222	SRR6145369
Eudigraphis taiwanensis	Transcriptome	Fernández *et al.*[7]	IZ-128912	SRR3458640
Glomeris marginata	Transcriptome	Fernández *et al.*[7]	IZ-43690	SRR3233211
Cyliosoma sp.	Transcriptome	Fernández *et al.*[7]	IZ-44064	SRR3458641
Brachycybe sp.	Transcriptome	Brewer & Bond[6]	—	SRX326776
Narceus americanus	Transcriptome	Fernández *et al.*[7]	IZ-44069	SRR3233222
Outgroups				
Limulus polyphemus	Transcriptome	Sharma *et al.*[37]	IZ:29738	SRX450966
Liphistius malayanus	Transcriptome	Sharma *et al.*[37]	IZ-29742	SRX450965
Centruroides vittatus	Transcriptome	Sharma *et al.*[37]	IZ-49754	SRX451012
Daphnia pulex	Genome		—	
Folsomia candida	Genome	Faddeeva-Vakhrusheva[38]	—	
Drosophila melanogaster	Genome		—	

Table 1. List of taxa included in the present study. Location, Museum of Comparative Zoology (MCZ) and SRA accession numbers are indicated.

generated from organisms targeted for their instability or lack of representation in prior analyses: a pauropod (*Pauropus huxleyi*) and a symphylan from the family Scolopendrellidae (Scutigerellidae was already represented in earlier studies). Information on sampling localities and accession numbers in the Sequence Read Archive database for each transcriptome can be found in Table 1. The remaining 12 myriapods from Brewer and Bond[6] and our own published data[7] were available from the Sequence Read Archive (SRA). The following taxa were included as outgroups: a crustacean (*Daphnia pulex*), two hexapods (*Drosophila melanogaster*, *Folsomia candida*), and three chelicerates (*Limulus polyphemus*, *Liphistius malayanus* and *Centruroides vittatus*). The new sequenced cDNA libraries are accessioned in SRA (Table 1). Tissue preservation and RNA sequencing are as described in Fernández *et al.*[8]. All molecular data included in this study were sequenced with the Illumina HiSeq 2500 platform.

Phylogenomic analyses. Single copy genes in arthropods were identified in our data sets with BUSCO v1.1[24] based on hidden Markov model profiles. The homologous genes detected were screened to identify multiple hits (i.e., paralogues). Only one homologue per BUSCO single copy gene was selected in each case, assuming that they were single copy in our samples as well, and therefore orthologs. The genes were parsed from each sample and combined into individual files (i.e., one file per gene) with custom python scripts. Alignment, trimming and concatenation were done as in Fernández *et al.*[7]. As the selection of outgroups may be critical in resolving myriapod relationships we constructed three matrices with different outgroup composition: matrix 1 (300 genes, 49,576 amino acids), including all outgroups (i.e., chelicerates, hexapods and crustaceans); matrix 2 (same as matrix 1, only with chelicerate outgroups); and matrix 3 (299 genes, 61,611 amino acids, only with pancrustacean outgroups). All matrices are provided as Suppl. Mat. In all cases, we selected a high level of gene occupancy to ensure the selection of a relatively large amount of genomic information while minimizing missing data and computational burden (75% gene occupancy in matrices 1 and 2 and 88% in matrix 3). Bayesian analyses were conducted with PhyloBayes MPI 1.7a[25] selecting the site-heterogeneous CAT-GTR model of amino acid substitution[26]. Two independent Markov chain Monte Carlo (MCMC) chains were run for 5000–10,000 cycles. The initial 25% of trees sampled in each MCMC run prior to convergence (judged when maximum bipartition discrepancies across chains were <0.1) were discarded as burn-in. Convergence of chains was assessed both at the level of the bipartition frequencies (with the command bpcomp) and the summary variables displayed in the trace files (with the command tracecomp). We considered that convergence was achieved when (i) the maximum difference of the frequency of all the bipartitions observed in the chains was <0.1, and (ii) when the maximum discrepancy observed for each column of the trace file was <0.1 and the minimum effective size of 100. A 50% majority-rule consensus tree was then computed from the remaining trees. In order to further test for the effect of heterotachy

Fossil taxon	Age (My)	Position
Rehbachiella kinnekullensis	497	crown Branchiopoda + Hexapoda
Rhyniella praecursor	405	crown Hexapoda
Casiogrammus ichythyeros	426.9	crown Diplopoda
Devonobius delta	382.7	crown Pleurostigmophora
Mazoscolopendra richardsoni	309.9	crown Amalpighiata
Gaspestria genselorum	392.1	crown Helminthomorpha
Proscorpius osborni	416	crown Arachnopulmonata
Crussolum sp.	408	crown Chilopoda

Table 2. List of fossil taxa selected for the molecular dating analyses. Their age and placement are indicated.

and heterogeneous substitution rates, the matrices were also analysed in PhyML v.3.0.3 implementing the integrated length (IL) approach[27,28]. In this analysis, the starting tree was set to the optimal parsimony tree and the FreeRate model[29] was selected. Congruence between the different topologies was visualized with DensiTree v2.2.5[30].

Molecular Dating. Divergence times for myriapods were estimated through molecular dating, constrained by the position of critical fossils using a morphological data set. Six Palaeozoic and Mesozoic myriapod fossils (three centipedes and three millipedes; Table 2) were included in our morphological data set of 187 characters. One fossil diplopod used for coding, *Cowiedesmus eroticopodus*, has since been redated as Early Devonian rather than mid Silurian[31]; another Silurian diplopod, *Casiogrammus ichthyeros*, replaces it as the earliest minimum age for crown-group Diplopoda and is coded as well.We also included three fossil outgroups: the crustacean *Rehbachiella kinnekullensis*, the scorpion *Proscorpius osborni*, and the collembolan *Rhyniella praecursor*. The matrix is available as Morphobank project P2762 (http://morphobank.org/permalink/P2762) and is provided as Supp. Mat. Multistate characters were scored as non-additive except for characters 57, 68, 81, 95 and 102, which were additive. The morphological data set was analysed under parsimony with TNT[32]. Traditional heuristic searches with 10,000 stepwise addition sequences resulted in 105 trees of 257 steps. Consistency Index 0.84, Retention Index 0.83 (Fig. 2e). No shorter trees were found using New Technology search strategies in TNT. Absolute dates follow the International Chronostratigraphic Chart v 2015/01. Justifications for age assignments of the fossils (Table 2) follow Wolfe *et al.*[33]. Divergence dates were estimated using the Bayesian relaxed molecular clock approach as implemented in PhyloBayes v.3.3[25]. Both an auto-correlated and uncorrelated relaxed clock model were applied to our dataset. The calibration constraints specified above were used with soft bounds[34] under a birth-death prior in PhyloBayes. Two independent MCMC chains were run for 5000–7,000 cycles, sampling posterior rates and dates every 10 cycles. The initial 25% were discarded as burn-in. Posterior estimates of divergence dates were then computed from the remaining samples of each chain.

References

1. Rota-Stabelli, O. *et al*. A congruent solution to arthropod phylogeny: phylogenomics, microRNAs and morphology support monophyletic Mandibulata. *Proc. R. Soc. B. Biol. Sci.* **278**, 298–306 (2011).
2. Rota-Stabelli, O., Daley, A. C. & Pisani, D. Molecular timetrees reveal a Cambrian colonization of land and a new scenario for ecdysozoan evolution. *Curr. Biol.* **23**, 392–398 (2013).
3. Lozano-Fernandez, J. *et al*. Molecular palaeobiological exploration of arthropod terrestrialisation. *Philos. Trans. Roy. Soc. B. Biol. Sci.* **371**, 20150133 (2016).
4. Misof, B. *et al*. Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**, 763–767 (2014).
5. Schwentner, M., Combosch, D. J., Nelson, J. P. & Giribet, G. A phylogenomic solution to the origin of insects by resolving crustacean-hexapod relationships. *Curr. Biol.* **27**, 1818–1824 (2017).
6. Brewer, M. S. & Bond, J. E. Ordinal-level phylogenomics of the arthropod class Diplopoda (millipedes) based on an analysis of 221 nuclear protein-coding loci generated using next-generation sequence analyses. *PLoS One* **8**, e79935 (2013).
7. Fernández, R., Edgecombe, G. D. & Giribet, G. Exploring phylogenetic relationships within Myriapoda and the effects of matrix composition and occupancy on phylogenomic reconstruction. *Syst. Biol.* **65**, 871–889 (2016).
8. Fernández, R. *et al*. Evaluating topological conflict in centipede phylogeny using transcriptomic data sets. *Mol. Biol. Evol.* **31**, 1500–1513 (2014).
9. Regier, J. C. *et al*. Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* **463**, 1079–1083 (2010).
10. Regier, J. C. & Zwick, A. Sources of signal in 62 protein-coding nuclear genes for higher-level phylogenetics of arthropods. *PLoS One* **6**, e23408 (2011).
11. Zwick, A., Regier, J. C. & Zwickl, D. J. Resolving discrepancy between nucleotides and amino acids in deep-level arthropod phylogenomics: Differentiating serine codons in 21-amino-acid models. *PLoS One* **7**, e47450 (2012).
12. Gai, Y.-H., Song, D.-X., Sun, H.-Y. & Zhou, K.-Y. Myriapod monophyly and relationships among myriapod classes based on nearly complete 28S and 18S rDNA sequences. *Zool. Sci.* **23**, 1101–1108 (2006).
13. Dong, Y. *et al*. The complete mitochondrial genome of *Pauropus longiramus* (Myriapoda: Pauropoda): Implications on early diversification of the myriapods revealed from comparative analysis. *Gene* **505**, 57–65 (2012).
14. Boudreaux, H. B. Phylogeny of the arthropod classes. In *Arthropod phylogeny*, *with special reference to insects*, pp. 82–122. Malabar, Florida: Robert E. Krieger Publishing Company (1987).
15. Pocock, R. I. On the classification of the tracheate Arthropoda. *Zool. Anz.* **16**, 271–275 (1983).
16. Tiegs, O. W. The development and affinities of the Pauropoda, based on a study of *Pauropus silvaticus*. *Quarterly. Journal of Microscopical Sciences* **88**, 275–336 (1947).
17. Rehm, P., Meusemann, K., Borner, J., Misof, B. & Burmester, T. Phylogenetic position of Myriapoda revealed by 454 transcriptome sequencing. *Mol. Phylogenet. Evol.* **77**, 25–33 (2014).

18. Miyazawa, H., Ueda, C., Yahata, K. & Su, Z. H. Molecular phylogeny of Myriapoda provides insights into evolutionary patterns of the mode in post-embryonic development. *Sci. Rep.* **4**, 4127 (2014).
19. Johnson, E. W., Briggs, D. E. G., Suthren, R. J., Wright, J. L. & Tunnicliff, S. P. Non-marine arthropod traces from the subaerial Ordovician Borrowdale Volcanic Group, English Lake District. *Geological Magazine* **131**, 395–406 (2014).
20. Wilson, H. M. Juliformian millipedes from the Lower Devonian of Euramerica: implications for the timing of millipede cladogenesis in the Paleozoic. *J. Paleontol.* **80**, 638–649 (2006).
21. MacNaughton, R. B. *et al*. First steps on land: Arthropod trackways in Cambrian-Ordovician Eolian sandstone, southeastern Ontario, Canada. *Geology* **30**, 391–394 (2002).
22. Collette, J. H., Gass, K. C. & Hagadorn, J. W. *Protichnites eremita* unshelled? Experimental model-based neoichnology and new evidence for a euthycarcinoid affinity for this ichnospecies. *J. Paleontol.* **86**, 442–454 (2012).
23. Blanke, A. & Wesener, T. Revival of forgotten characters and modern imaging techniques help to produce a robust phylogeny of the Diplopoda (Arthropoda, Myriapoda). *Arthropod Struct. Dev.* **43**, 63–75 (2014).
24. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
25. Lartillot, N., Rodrigue, N., Stubbs, D. & Richer, J. PhyloBayes MPI: Phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Syst. Biol.* **62**, 611–615 (2013).
26. Lartillot, N. & Philippe, H. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol. Biol. Evol.* **21**, 1095–1109 (2004).
27. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704 (2003).
28. Guindon, S. From trajectories to averages: an improved description of the heterogeneity of substitution rates along lineages. *Syst. Biol.* **62**, 22–34 (2013).
29. Soubrier, J. *et al*. The influence of rate heterogeneity among sites on the time dependence of molecular rates. *Mol. Biol. Evol.* **29**, 3345–3358 (2012).
30. Bouckaert, R. R. & Heled, J. DensiTree 2: Seeing trees through the forest. *bioRxiv* (2014)
31. Suarez, S. E., Brookfield, M. E., Catlos, E. J. & Stöckli, D. F. A U-Pb zircon age constraint on the oldest-recorded air-breathing land animal. *PLoS One* **12**, e0179262 (2017).
32. Goloboff, P. A. & Catalano, S. A. TNT version 1.5, including a full implementation of phylogenetic morphometrics. *Cladistics* **32**, 221–238 (2016).
33. Wolfe, J. M., Daley, A. C., Legg, D. A. & Edgecombe, G. D. Fossil calibrations for the arthropod Tree of Life. *Earth-Sci. Rev.* **160**, 43–110 (2016).
34. Yang, Z. H. & Rannala, B. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* **23**, 212–226 (2006).
35. Chipman, A. D. *et al*. The first myriapod genome sequence reveals conservative arthropod gene content and genome organisation in the centipede *Strigamia maritima*. *PLoS Biol.* **12**, e1002005 (2014).
36. Stoev, P. *et al*. *Eupolybothrus cavernicolus* Komerički & Stoev sp. n. (Chilopoda: Lithobiomorpha: Lithobiidae): the first eukaryotic species description combining transcriptomic. *DNA barcoding and micro-CT imaging data. Biodiversity data journal* **1**, e1013 (2013).
37. Sharma, P. P. *et al*. Phylogenomic interrogation of Arachnida reveals systemic conflicts in phylogenetic signal. *Mol. Biol. Evol.* **31**, 2963–2984 (2014).
38. Faddeeva-Vakhrusheva, A. *et al*. Coping with living in the soil: the genome of the parthenogenetic springtail *Folsomia candida*. *BMC Genomics* **18**, 493 (2017).

Acknowledgements

Miquel Arnedo, Ligia Benavides, Brian Colby and Julia Cosgrove assisted with pauropod collection. This study was mainly supported by internal funds from the Museum of Comparative Zoology, Harvard University and by NSF grant DEB-1457539 to GG, which funded RF. Three reviewers provided comments that help to clarify some aspects of this article.

Author Contributions

All authors designed the study. R.F. and G.G. collected the samples. R.F. executed molecular work and data analysis. G.D.E. coded the morphological data. All authors wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-017-18562-w.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.